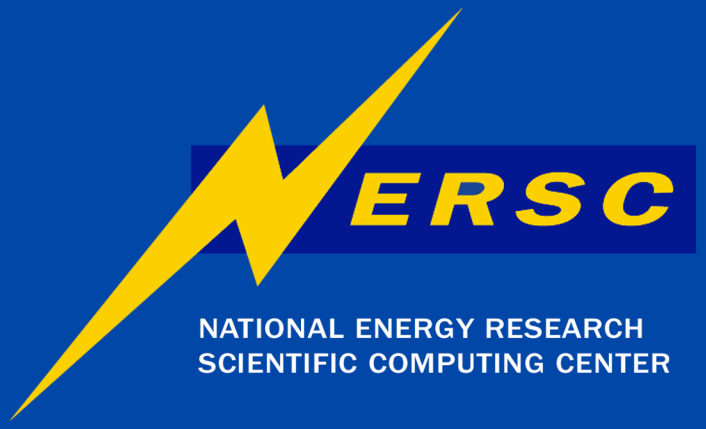


Capability Panel Discussion

Horst D. Simon
Associate Lab. Director, Computing Sciences,
Berkeley Lab
hdsimon@lbl.gov

SOS 10 Workshop
March 10, 2006

Capability Panel

- Horst D. Simon, NERSC, Berkeley Lab
- Ruud Haring, IBM
- John Morrison, LANL
- Marie-Christine Sawley, CSCS
- Jim Tompkins, Sandia
- Al Geist, ORNL

"Capability" as defined by NRC

- The largest supercomputers are used for capability or turnaround computing where the maximum processing power is applied to a single problem.
  - solve a single problem in a shorter period of time
  - enables the solution of problems that cannot otherwise be solved in a reasonable period of time
  - enables the solution of problems with real-time constraints
  - figure of merit is time to solution
  - designed to offer the best possible capability, even at the expense of increased cost per sustained performance

"Capacity" as defined by NRC

- Smaller or cheaper systems are used for capacity computing, where smaller problems are solved.
  - can be used to enable parametric studies or to explore design alternatives
  - often needed to prepare for more expensive runs on capability systems
  - run several jobs simultaneously
  - figure of merit is sustained performance per unit cost
  - designed to offer a less aggressive reduction in time to solution but at a lower cost per sustained performance

Questions for Panel

1. Briefly describe the capability resources at your site.
2. By example describe one or two applications, where your unique capability platform was critical in providing a solution.
3. Do you agree with the above distinction between capability and capacity? If not, how would you define these terms?
4. Is the distinction between capability and capacity useful?
5. What metrics do you use to measure "capability"?
6. How could we as a community improve these metrics?

Neil said: “say something about TOP500”

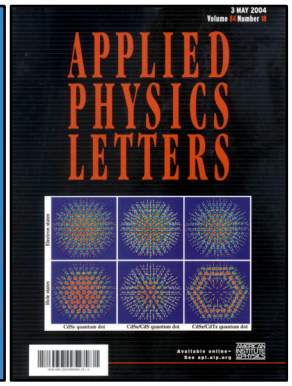
“Whoever buys a system to reach a certain rank on the TOP500 list is an idiot.”

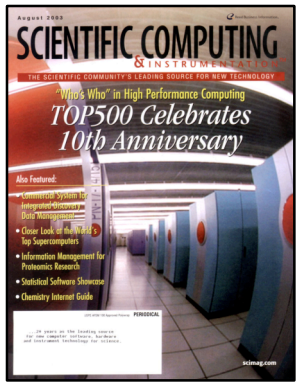
TOP500 does not define capability computing

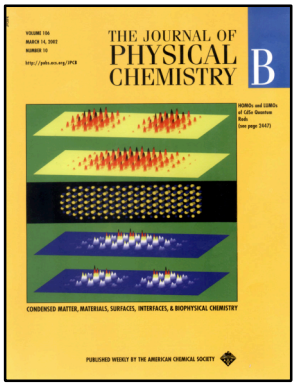
TOP500 Restrictions

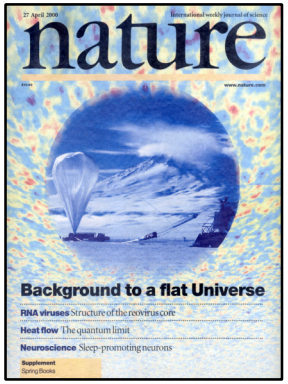
- The authors reserve the right to independently verify submitted LINPACK results, and exclude systems from the list which are not valid or not general purpose in nature.
- As an example, a system that is built with special purpose hardware accelerators to perform well on just the LINPACK benchmark, but that has limited value for real applications could be excluded from the list.
- The TOP500 authors also reserve the right to exclude any system from the TOP500 list that is not yet installed at an end user site, or that is not yet ready for general application use.

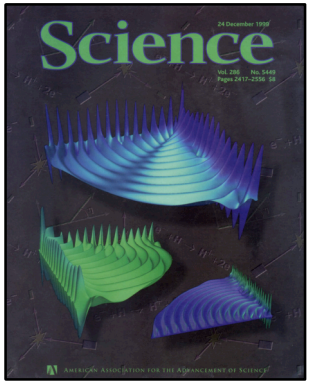
Science-Driven Computing

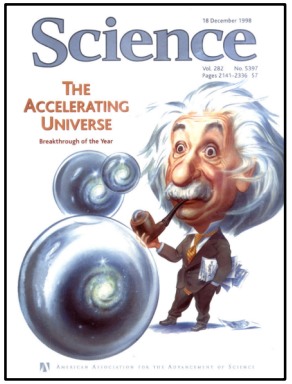






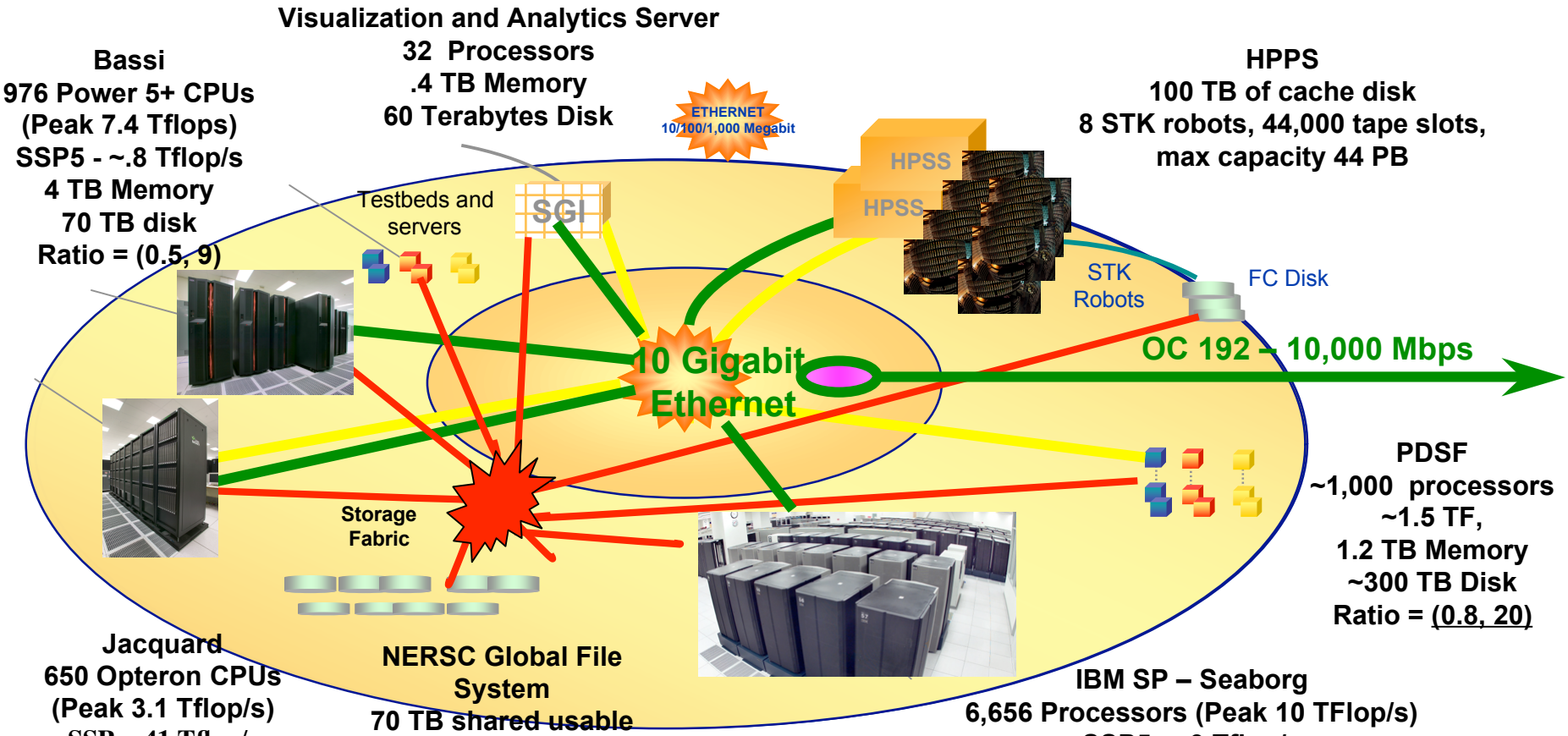






NERSC's role is to enable new science

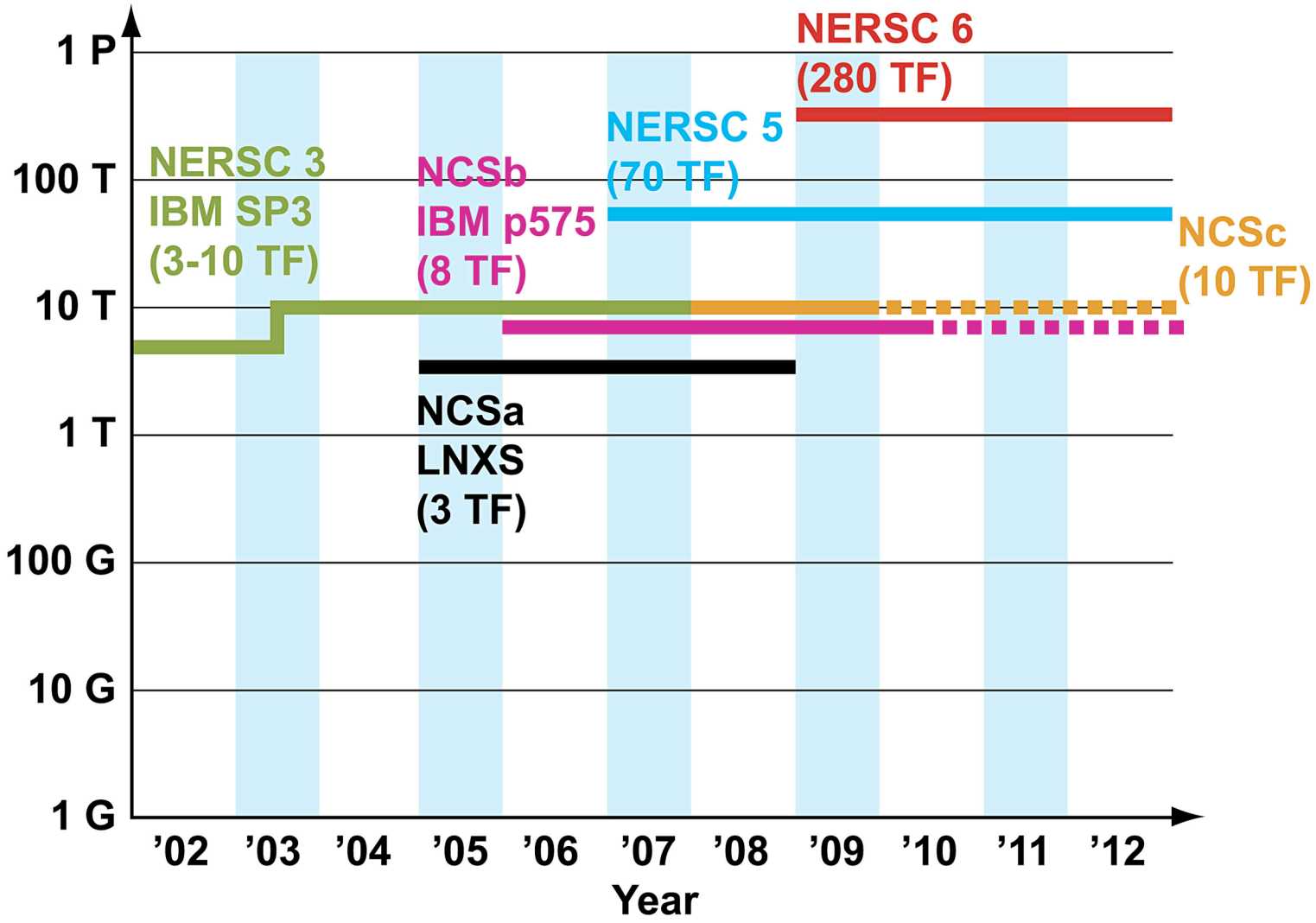
NERSC Configuration
January 2006



Ratio = (RAM Bytes per Flop, Disk Bytes per Flop)

Office of Science
U.S. DEPARTMENT OF ENERGY

NERSC Systems Roadmap



Office of Science
U.S. DEPARTMENT OF ENERGY

NERSC Capability is more than cycles

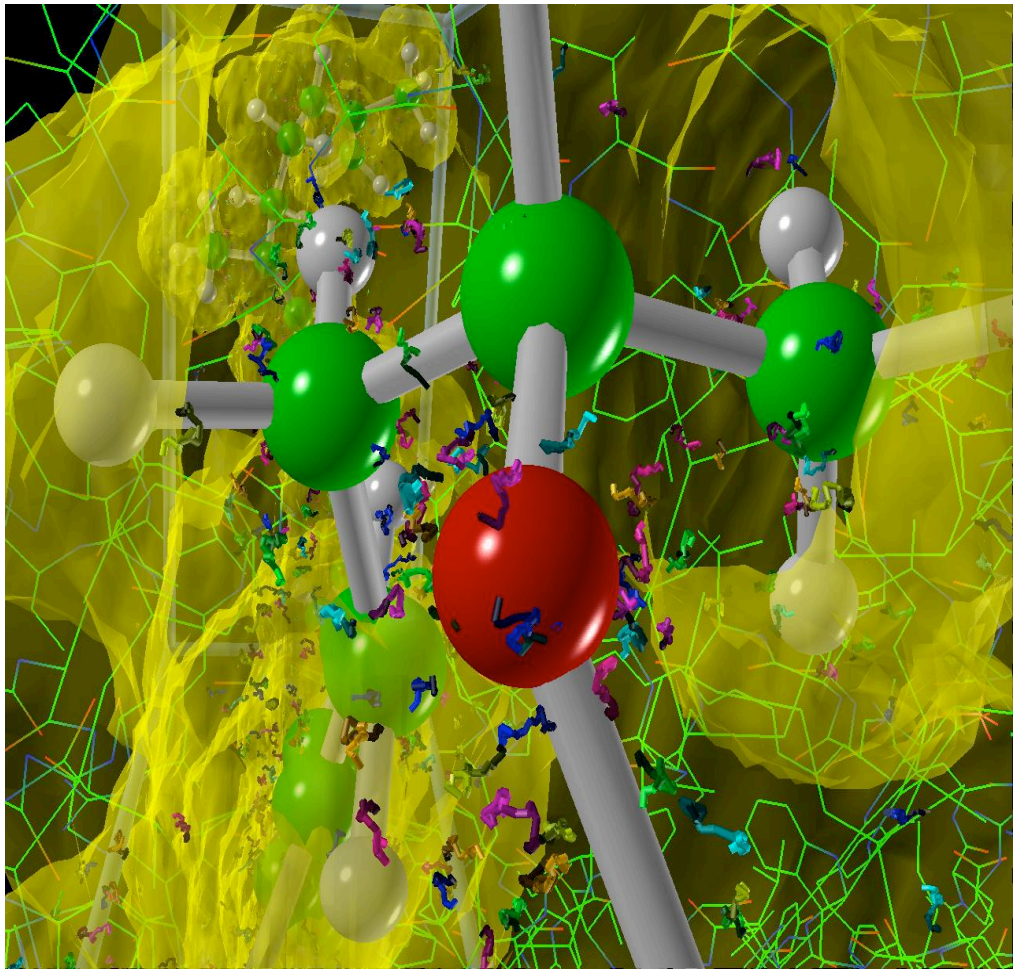
Photosynthesis INCITE Project

PI: William Lester, UC Berkeley

- 2M hours provided
- MPI tuning: 15-40% less MPI time
- Quantum Monte Carlo load balancing: 256 to 4,096 procs
- More efficient algorithm for random walk procedure
- Wrote parallel HDF5 I/O layer
- Used AVS/Express to visualize molecules and electron trajectories
- "Visualization has provided us with modes of presenting our work beyond our wildest imagination"
- "We have benefited enormously from the support of NERSC staff"

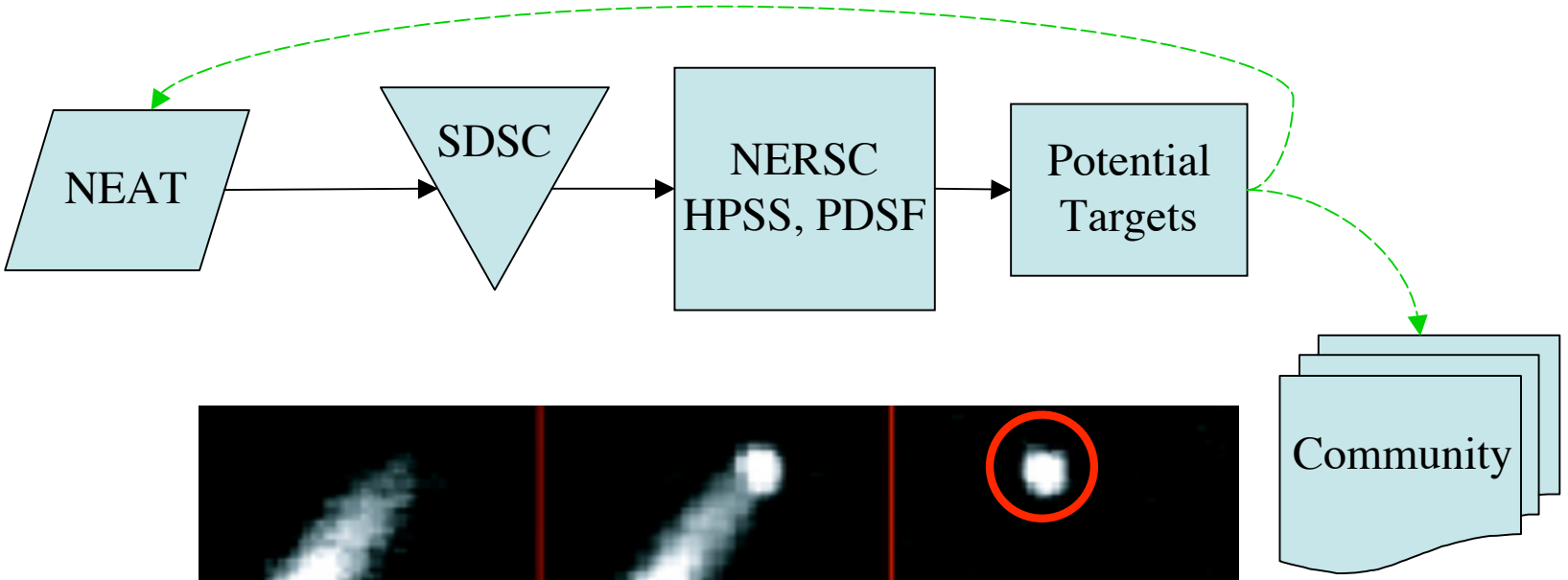
Largest QMC calculation today;

Enabled comparison to local density functional theory approach



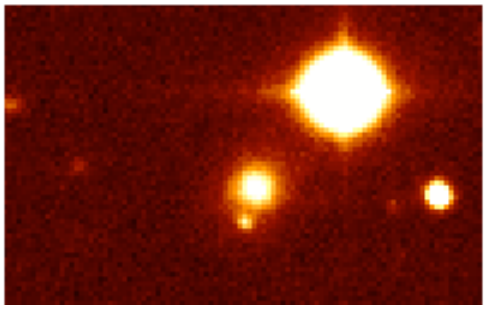
Feedback Loop to Experiment

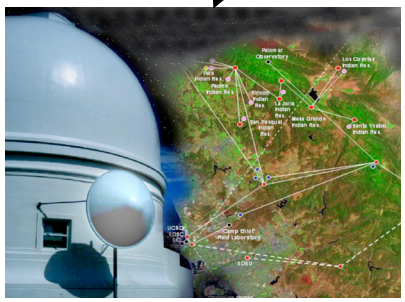
- **Nearby Supernova Factory data analysis pipeline:**

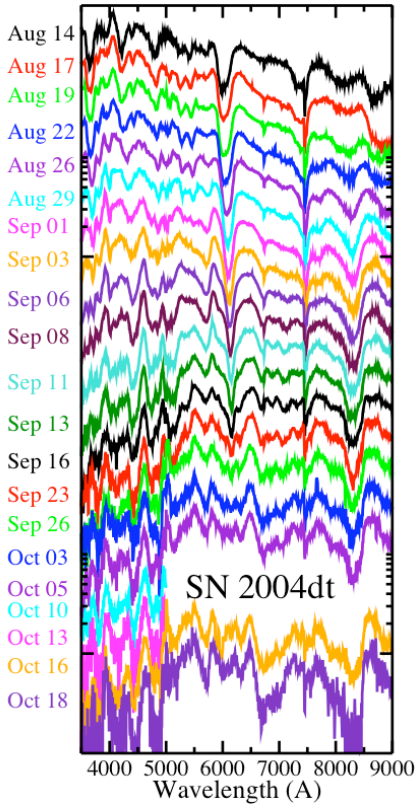


NERSC capability is providing the integrated workflow

- Goal: Find and examine in detail up to 300 nearby Type Ia supernovae
  - Detailed, high statistics sample against which distant SN can be compared
- Discovered 50+ confirmed supernovae in 2005
- Previously, there were fewer than 50 cosmologically useful nearby supernovae
- Processed 7 million images to date, archived as 20 TB of compressed data.
- Spectral database larger & better than entire body published since 1937
- NERSC contributions:
  - high-speed data link
  - automated analysis of spectra
  - machine-learning applied to candidate SNe
  - NERSC's ability to store and process 50 gigabytes of data every night
  - long- and short-term scheduler







Office of Science
U.S. DEPARTMENT OF ENERGY

Capability versus Capacity

- The capability/capacity distinction is not useful, and potentially harmful
- We need to distinguish at least three dimensions
  - System/Architecture: special purpose/customized vs. commodity
  - Metric: time to solution vs. sustained performance/unit cost
  - Policy: system for elite few vs. system for large user workload
- The current use of capability and capacity groups these three dimension together into only one

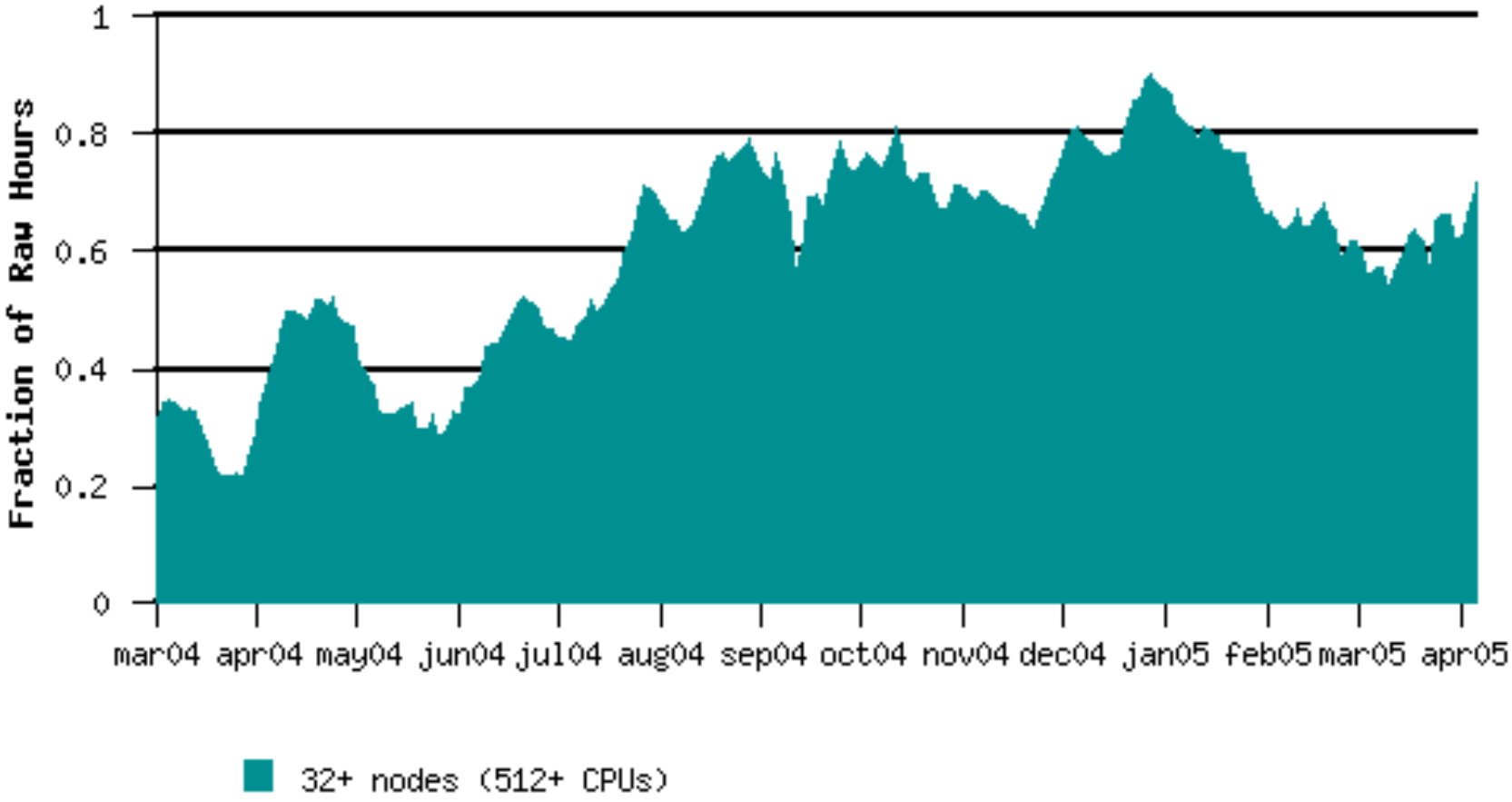
Capability versus Capacity

- Large Facilities such as NERSC have had to justify their existence by providing capabilities that are not likely to be available in other situations (labs, universities,etc)
- Two Issues
  - There is continued pressure towards capacity computing by allocation
    - Allocate more users, give many small pieces, don't make the hard priorities
  - There is continued pressure towards capability computing by political justification

Metrics

- DOE committed to OMB that "50% of computational time will be used by jobs 1/8th [512 CPUs] or more of the total possible resource"
  - Yet, only 40% of the allocation is provided to jobs that have shown the ability to run at this level.
- Metrics put centers like NERSC in a difficult position.
  - Focus on number of CPUs used simultaneously ignores memory, I/O, etc.
- NERSC has aggressive scheduling and charging benefits for large jobs, and a large job reimbursement program to meet the metric

NERSC Focus Is on Large-Scale Computational Science



Capability or Capacity?

- All complex science has a range of types and sizes for jobs
  - Does a 1,000 year climate run with the latest code and physics need capability computing? (208 CPUs running continuously for 1 year)
  - A colliding black hole simulation used more than 3 TB of memory and had time step files between 2-5 TB in size. Does it matter if it were to run on "only" 256 processors?

Capability Definition

- Capability computing uses large amounts of integrated resources over time
  - Resources are CPU time, memory, storage, bandwidth, software, expertise
  - "Over time" must mean that these resources are predictably and consistently available
- A capability center provides capability resources that enable unique computational simulations and analysis that cannot be easily carried out anywhere else
- The metric is the quality of science produced